PLATTFORM ftEval
Forschungs- und Technologieevaluierung

Karriereverläufe

Nr. 36
Juni 2011

Grit Laudel and Jochen Gläser

Academic careers and how to find research excellence

Introduction

It is difficult to find another work process as suffused with evaluations as research. Research uses scarce resources and produces results other researchers need to use and thus must be able to rely on. Consequently, researchers are evaluated when they start a project (trough the peer review of project grants), while they conduct it (through collaborations and informal communication), when they publish their results (through the peer review of their manuscripts), and when they apply for new jobs (through the peer review of applicants' research performance). On average, a researcher can be assumed to undergo an evaluation at least every two months.

These constant evaluations are esoteric and difficult to assess outside the scientific community that constitutes a researcher's primary work context. Judgments are made implicitly and informally, or are formulated in a way that their meaning is fully accessible only to members of the same community. While there are events that are regarded as 'success' or 'failure' by outsiders, those events are comparatively rare and are still open to interpretation. The constant evaluation of researchers is an inseparable part of an esoteric work process and thus difficult to understand.

These evaluations have recently been topped by a new layer . Driven by the twin moves to more autonomy and more accountability, research organizations feel the need to improve their research performance. Improving research performance depends on knowing it, and knowing research performance requires evaluations. Thus, organizations have been establishing their own research performance evaluations of subunits and individuals. They ask for numbers of publications, indicators for the quality of publications, numbers and amounts of research grants, prizes and awards, and so on. Researchers begin to adapt to this new interest by proactively including the required information in their CVs, grant applications, and web pages.

But do the indicators that are commonly used tell us anything about research performance at all? What aspects of performance, if any, do they measure? Can they, via researchers adapting to them, distort research and research performance? In this contribution, we discuss the uses of indicators that can be derived from researchers' careers. We demonstrate that most of the simple indicators are close to useless, and understanding researchers' performance requires a thorough study of their career.

1. Which career?

Both our everyday experience and organizational theory consider careers as happening within, and in some cases also between, organizations. Careers are seen as a sequence of moves between positions, and the major outcome of such moves is usually regarded as an increase in responsibilities, power, and income.

Researchers are different in that they have three careers at once. The first career is their research career, i.e. the sequence of interlinked projects they work on. These projects often build on each other, and each of them contributes to an academic's perspective on what to do next as well as to the academic's abilities to solve particular problems. They thus constitute a series of path dependent steps that is characteristic of careers (Barley 1989).

Closely linked to this research career is an researcher's community career. By 'community career' we mean the move of researchers between status positions in the social context that is crucial for their research (and much more important than their organization) - their scientific community. The scientific community is the social collective that produces scientific knowledge. Researchers derive their individual plans from their perception of the community's shared knowledge, and offer their own contributions to that knowledge in publications. The scientific community is therefore an often invisible but nevertheless very powerful social context for researchers (Gläser 2006). Academics go through a status career in their scientific community that comprises distinct positions. We use ideas about professional careers (Dalton et al. 1977) to distinguish four stages of an academic's community career (Gläser 2001: 703):

1. An *apprentice* learns to conduct research while working under the direction of others. PhD students are apprentices, in some cases this stage may extend to early postdoctoral phases.
2. A *colleague* conducts independent research, i.e. autonomously decides on problems to solve, on approaches to problem solving, and on ways to communicate results to the scientific community.
3. A *master* additionally acts as a teacher of apprentices.
4. A member of the *elite* additionally shapes the direction of the knowledge production of their community by exceptional research contributions, by taking on influential roles in professional organizations, funding organizations, and other interfaces between science and society, and by informallysteering the community through elite networks (Laudel 2005a: 390).

Most researchers – but not all of them - will reach the position of a colleague. Many will also reach the level at which they supervise PhD students, i.e. the master stage. The last stage of an elite member is of course achieved by only a handful of researchers.

Finally, researchers also have a 'conventional' organisational career. Most current research would be impossible without organizations providing researchers with positions (and thus an income) and resources for their research. By moving between jobs offered by these organisations researchers go through an organisational career whose stages are linked to specific performance expectations and research opportunities.

However, organizations can affect the directions and quality of research only indirectly. Decisions on research content are made by individual researchers in the context of their community. Organisations can therefore be best thought of as merely 'hosting' researchers who 'rent' places in organisations that provide them with income and resources. They use these opportunities to contribute to the knowledge of their community and 'pay' the organization with their reputation and contributions to other organizational tasks such as teaching (Clark 1983: 28-34; Sørensen 1992: 94-96). Another peculiarity of research organizations is that they offer little in terms of intraorganizational careers. Owing to the specialization of research communities, there are very few steps on the intraorganizational career ladder for researchers. It is impossible for an assistant professor in ancient history to move to the position of an associate professor in quantum physics that recently has become vacant at the same university. This is why there are only very few intraorganizational moves between positions, and why moves between organizations are relatively frequent.

When we want to use 'career indicators' for assessing the quality of research and researchers, this complex interdependent system of three careers confronts us with a fundamental problem. Nobody outside a researchers' scientific community is able to directly judge the quality of research. Therefore, anybody outside the community – i.e. anybody who doesn't understand the content of the research undertaken - is forced to resort to secondary indicators. Secondary indicators are those that describe properties of the research and the researcher which are hoped to reflect quality.

Unfortunately, these indicators offer only partial insights into research quality at best, and are usually unreliable when applied to individual researchers and their research. In the following two sections we discuss the most severe problems of commonly used indicators of individual research performance and point out what cannot be learned and what can be learned from indicators describing academic careers.

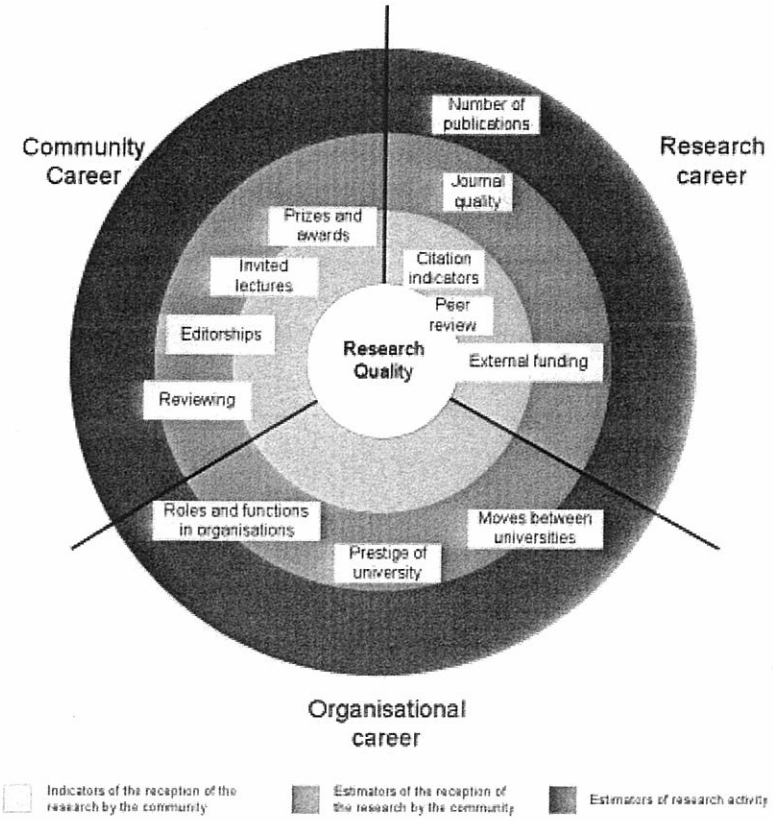
2. Commonly used indicators and their pitfalls

All three careers of a researcher offer indicators of research performance. Owing to the specific links between the careers, these indicators differ in their distance from the actual research quality. A researcher's status career in the scientific community depends on the quality of results produced during the research career, and the organizational career depends on both. Thus, using career indicators depends not only on the technical difficulties of measuring research quality at all but also on the conceptual difficulty of linking an indicator to what is intended to measure, namely the quality of research or of the researcher.

We sort our discussion of commonly used indicators along the conceptual dimension, i.e. the distance respectively the number of translations required between research quality and the indicators. *Diagram 1* orders career-based indicators according to the researcher's career they are linked to and according to their distance from research quality. This distance is the major reason for the limited validity of the indicators.

Diagram 1

Indicators based on a researcher's three careers and their relative distance from research quality.



Indicators derived from the research career

As we already stated in the previous section, only a researcher's peers – researchers working in the same field – can directly assess the content of research and therefore the quality of results. We placed 'peer review' in the diagram closest to research quality in order to highlight this fact. Since peer review is also based on colleagues' opinions about a piece of research, we placed it across the 'research quality' and 'indicators of reception' zones.

Citation indicators and indicators based on external funding touch on research quality because they contain an element of peer review – researchers assessing the utility of a published piece of research or reviewers assessing the promise of a research idea [1]. However, in both cases the peer review is slightly 'off the mark' of research quality. In the case of the citation-based indicators the peer reviews are individual decisions on relevance and usefulness rather than quality. This is why bibliometricians have agreed on the formula that citations measure one important aspect of research quality, namely international impact (e.g. Van Raan 1996: 404). In the case of external funding, the peer review addresses the originality of a research idea but also other criteria such as feasibility, relevance, and adherence to other criteria of funding programmes.

Both citation and funding indicators are also burdened with enormous technical problems that make it difficult if not impossible to use them at the level of the individual researcher. The first major problem is statistical validity. The indicators have to be thought of as random variables, which means that enough instances must be provided for them to be valid. It has been estimated that a basis of ten to twenty publications per year is sufficient (van Raan 2000: 307-309), which means that bibliometric indicators are not applicable to most individual researchers and are sometimes problematic even at the levels of research groups. In the light of this estimate, recent attempts to use citation-based indicators for the evaluation of individuals (Costas and Bordon 2005; Costas et al. 2010) must be considered both technically and ethically problematic. The same applies to external funding, which is, however, less systematically used at the individual level. Assessing the external funding a researcher has acquired is more of an interpretive exercise, with number of grants or sums not being used as hard information in comparisons.

Both kinds of indicators also have in common that they are field-specific. Citation indicators cannot be applied in most fields because the available databases cover only a certain proportion of journal articles from each field. The lower the proportion of journal articles in the publications of a field and the lower the proportion of the field's journals covered by the databases, the lower the validity of citation indicators [2]. Furthermore, numbers of citations as well as the frequencies and amounts of grants starkly vary between fields of research and countries. The sizes, research practices and publication practices of fields of research affect the numbers of citations a publication will typically receive. A field's resource intensity and form of institutionalisation at universities will affect the numbers and sizes of grants. The country a researcher works in will further modify these characteristic averages of fields because they also depend on the publication language (because of the English language bias of citation databases) and on the proportion of research funding that is allocated in form of competitive grants.

[1] Commonly used citation indicators include the absolute number of citations, citations per publication, and the Hirsch index – a number h that signifies the number of a researcher's publications that have been cited h times or more. The Hirsch index is believed to reflect the total impact of a researcher'scontributions so far (Hirsch 2005, Bornmann and Daniel 2009). Numbers of competitive, peer reviewed grants or the amount of competitive grant funding are used as indicators of external funding.

[2] In a study of Australia's publication output, Butler and Visser (2006: 329) found that 90% of publications in chemistry but only 6% of law publications were covered by the Web of Science.

In the case of most citation indicators, normalisation techniques have been developed to solve the problem of field specificity. While the same is possible in principle for external funding indicators, no serious attempt has been made so far. It seems doubtful that such a normalisation is possible at all at the individual level. Citation and funding indicators are also blind to the variation of individual research and publication styles. While the continuation of research is impossible without grants in the experimental sciences in countries with a high proportion of external funding, there are always fields in which research can be conducted either with or without external funding. External funding indicators tell something about research activity if not quality in fields of the first kind, but are heavily affected by individual styles of using or not using external funding in fields of the latter kind.

The case about styles has found sufficient backing from the first studies using citations as indicators. We illustrate the problem by confronting the Hirsch index with a distinction made by Cole and Cole (1967) in their analysis of physicists in the US (diagram 2). While the Hirsch Index would reflect research performance as we wish in the case of prolific and silent researches, both mass producers and perfectionists could end up with the same Hirsch index. Somebody who produces 400 publications, of which only four are cited four times and more, has a Hirsch index of four, just as somebody who produces four publications that are cited more than four hundred times each.

Diagram 2
The Hirsch index for four types of researchers (definition of types adopted from Cole and Cole 1967: 381-382)

		Numbers of citations	
		High	Low
Numbers of publications	High	Prolific high Hirsch index	Mass producers low Hirsch index
	Low	low Hirsch index	low Hirsch index

Apart from being inherently problematic at the individual level due to the low numbers involved, citation indicators are biased against non-Anglophone countries and may be biased against particular styles of research. Indicators of external funding have the additional problem that the success in receiving external grants depends by no means on the quality of the researcher alone. The extent to which researchers can rely on recurrent funding in their preparation of the grant application, the match between a researcher's interests and political priorities, the extent to which a proposal extends what a researcher has previously worked on and other factors have been shown to affect the success in grant funding (Laudel 2005b).

In the next zone of the research career indicators we find estimators of the reception by the community. Proponents of such indicators refer to the high costs of valid citation analyses, which need to be conducted by specialists, and to the fact that citation analyses cannot be conducted for all fields. As a shortcut, the journal impact factor or a journal rank is used. The impact factor is intended to predict how many citations an article published in a journal can be expected

to receive [3]. Since many journals don't receive an impact factor because they are not indexed by the Web of Science, several attempts have been made to rank journals according to their quality [4]. In both cases, the quality of a journal that is constructed as the average quality of previously published articles is used to assess the quality of a new article. This is obviously risky. The distribution of quality of articles is highly skewed, which means that the reputation of any journal is earned by few articles, while most of them gain little recognition. Again, the approach might work in the statistical aggregate but it is simply impossible to draw any conclusion from impact factors or journal ranks to the quality of a individual article published in that journal.

We conclude the discussion of indicators pertaining to the research career by briefly stating that the number of publications is not a valid indicator of research performance. The number of publications is simply an indicator of activity. A researcher who doesn't publish is unlikely to conduct research. Beyond that fact, numbers of publications offer little useful information. They are even dangerous because their use can incite mass production of useless publications.

Indicators derived from the community career

Indicators derived from the other careers are inevitably more distant from research quality than peer review, citation indicators, or external funding because they are based on summary judgments of the research or community career. The community career provides various indicators of esteem by the community which include prizes and awards, invited lectures, editorships of journals and book series, and invitations to review manuscripts or grant proposals. The logic underlying the use of these indicators is always the same: These events indicate that the community values the quality of research and the qualification of the researcher in question. This is why they are located in the second zone of *diagram 1* as indicators of reception, albeit in varying distance from research quality itself. The reasons why the validity of these indicators is problematic in each individual case in spite of the logic being generally sound are:

- The indicators are idiosyncratic and therefore not comparable in most cases.
- The rapid growth of the communication system and the increasing demand for reviewers puts pressure on the quality criteria applied in the selection of reviewers.
- In some scientific communities one can observe an internal stratification in which editorship, reviewing, and invited lectures occur on all different quality strata of the community.

These problems cannot be overcome by any kind of normalisation. Instead, a very careful analysis of individual cases is required in order to assess what an indicator of esteem actually means.

Indicators derived from the organisational career

If indicators of esteem are already more distant from research quality than the most valid indicators of research quality, than this is even more the case for indicators derived from the researcher's organisational career. These indicators refer to organisational roles that are assigned to researchers on the basis of their status in their scientific community, which in turn is based on their research. The logic underlying these indicators is that researchers compete for better organisational positions, and that these positions are allocated according to the reputation of researchers, which in turn derives from research quality.

[3] The impact factor is calculated by counting the number of citations received by articles in two subsequent years in the following year and dividing it by the number of citable items (articles, reviews, proceedings, or notes; not editorials or letters-to-the-editor) published in these two years. The impact factor for 2011 will be calculated by dividing all citations received in 2011 by articles published in 2009 and 2010 by the number of citable items published in 2009 and 2010.

[4] The ranking of journals is usually based on opinion polls among academics.

The indicator 'working at a prestigious organisation' is construed in a similar way as the journal indicators: the reputation previously accrued by the university's other researchers is used as an estimator of the quality of all its current members. *Table 1* demonstrates that such an indicator is close to useless by comparing the quality of research of two British universities as measured in the Research Assessment Exercise 2008. While the University of Oxford as a whole conducts better research in Psychiatry, Neuroscience and Clinical Psychology than the University of Sheffield, a randomly selected researcher from Sheffield could be one who conducts research that is as good as that of researchers at Oxford.

Table 1

Quality profiles of two universities' research in Psychiatry, Neuroscience and Clinical Psychology according to the Research Assessment Exercise 2008 (source: RAE 2008)

	PROPORTION OF RESEARCH THAT IS			
	World-leading	**Internationally excellent**	**Recognised internationally**	**Recognised internationally**
University of Oxford	15%	45%	30%	10%
University of Sheffield	10%	30%	50%	10%

The other indicators are at least as problematic as the prestige of the university. Moves between universities occur for a variety of reasons, many of which have nothing to do with research. Similarly, the logic according to which organisations assign positions is often not purely based on research quality. Organisations have other internal tasks and thus use more than one criterion to assign positions, roles, and functions.

The summary of this discussion of common career indicators is obviously pessimistic. The main reason for the impression that no valid indicators of individual research performance can be derived from a researcher's three careers is that the few indicators that somehow work at the aggregate level lose their power if applied to individuals. Each individual case we look at may be contrary to the trend that justifies the application of an indicator to groups or organisations. In order to learn something about an individual case, we need to study it in depth.

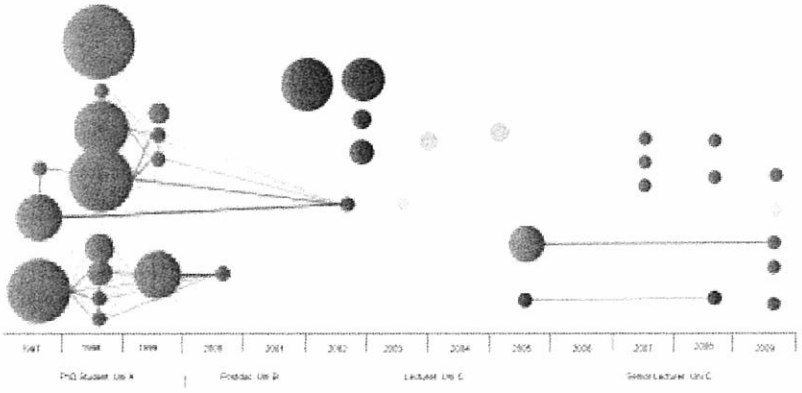
3. Process-oriented indicators

Process-oriented indicators reflect the dynamics of the research career rather than summarizing properties of the whole career or stages of it. They are both more difficult to use and more 'interesting' because they offer insights the traditional (and, as we have seen, questionable) indicators don't. As all indicators of research performance, process-oriented indicators require interpretation by peers and preferably a conversation with the researcher whose career they

describe. We base our discussion of such indicators on a visual representation of a researcher's research career and organizational career which we use in our interviews with academics *(diagram 3)*. The picture shows the publication history of a researcher and the organizational positions as they change over time [5]. We use it for discussing characteristic indicators of successful apprenticeship and colleague stages as well as the elite stage.

Diagram 3

Visual representation of a researchers research career (as reflected in thematically linked publications and citations of them) and organizational career.



Apprenticeships

In almost all careers we have investigated so far, successful researchers already had successful apprenticeships, i.e. PhD phases. This means that apprentices became visible to their community and produced results that were used by others. In terms of traditional indicators, this means that their PhD research was published and cited. As the many publications and citations indicate, the researcher whose careers are represented in diagram 3 had a very successful PhD phase. In the social sciences and the humanities, we have to apply different criteria. It is expected in most disciplines that the PhD results are published as a book (besides other forms of publishing results). The book is received by the community in form of book reviews (the mere existence of which is already an indicator of success) and citations to the book.

The numbers of publications and citations characterizing a successful PhD depend on many factors including the research field the researcher's country, the publication language, and the concrete research problem. The interpretation of these numbers is difficult because in many cases, the research problem has not been formulated by the PhD student but by the supervisor. Thus, part of the success of PhD theses depends on the quality of supervisors and their research groups. More generally speaking, results of PhD research in the sciences often are outcomes of collaborative efforts involving the supervisor and other lab members. Since most PhD students do not conduct independent research, we must be cautious in interpreting the PhD success. However, a high quality research environment is an important 'breeding ground' for successful careers and therefore.

Colleagues

An apprentice researcher's PhD phase triggers the start of new research trails that are thematically different from the

[5] For scientists, publications are downloaded from the Web of Science and clusters of thematically linked publications are identified by bibliographic coupling using Salton's cosine as measure of the strength of shared references. In the case of researchers from the social sciences and humanities, publications are obtained from CVs clustered on the basis of lexical couplin (shared keywords in titles). The sizes of circles indicate numbers of citations, colours indicate thematically different clusters (Gläser and Laudel 2009).

PhD work. Scientists use postdoctoral phases to broaden their thematic background and to learn more methods. They subsequently move to an academic position at a university and build their own research groups and laboratories. In many fields, this process is reflected in the positions on author lists. Most scientific research is collaborative, and in many fields there is an informal rule about the order of a journal article's author list. According to this rule, the researcher who conducted the experiments (in many cases a PhD student) becomes first author, the head of the laboratory (who often defined the problem and contributed theoretical background) is last author, and other collaborators (including postdocs to whom the supervision of the PhD might have been delegated) are listed in between. In fields where this rule applies, the progress of a researcher's career is reflected by the move of his or her name from first author to intermediate author to last author.

Elite

Elite researchers influence their scientific communities through their own research contributions as well as their formal roles and their informal decision-making in elite networks. The influential research contributions often consist of linking fields and ideas that their community thought of as separate. By creating such new interdisciplinary links, elite researchers often open up whole new areas of research opportunities for one or more communities.

Another characteristic of elite researchers is that they have long-term research programmes. One of our interviewees, a physicist, spent six years on the development of a new method, which required a new instrument. He then spent three years on the improvement of that method/instrument. During this time (after seven years of the development) he started using this instrument for studying biological objects. He plans to continue these investigations for another five years after the improvement of the instrument is finished. Thus, his personal research programme spans a time of at least 14 years.

Finally, elite researchers can be identified by their collaboration networks. Scientists don't just collaborate with anybody who can contribute the knowledge they need. They always try to find the best experts on this knowledge and to collaborate with them. Thus, collaboration networks are stratified according to the quality of collaborators [6]. This means that researchers who collaborate with the best groups in a field are likely to belong to the elite themselves.

Process-oriented indicators are even more difficult to obtain than the few valid traditional indicators of research performance, and equally difficult to interpret. However, a researcher's CV is a good source of information about most of the indicators we have discussed in this session. If any judgment is to be based on the indicators discussed in this section at all, this requires of course discussing the various indicators with the researcher in some depth. However, this is the best way to obtain information about a researcher's career anyway.

4. Conclusion: How can career indicators be used?

There is no easy way to judge the quality of research or of the researchers by indicators derived from a researcher's careers, and we don't have anything else. Most of the indicators are difficult to apply, lack statistical validity at the individual level, or are difficult to interpret. Our discussion can be summarized as follows: If

- many of the traditional indicators point in the same direction over a longer period of time,
- show extreme values (e.g. very many or no citations, very prestigious or no awards, very competitive or no external funding and so on), and

[6] There are of course exceptions to that rule, particularly due to path dependencies. This means just that as always, the individual case might deviate from the pattern.

– are interpreted in the relevant contexts of field, country, and language, then they can be used to identify the extremes of a researcher population, namely the elite and the academics who don't conduct research at all. For assessing the rest of the population, process-based indicators and in-depth analysis of careers appear to be necessary to do justice to individual cases.

Finally, we would like to warn against the aggregation of career indicators. We emphasized the three careers of a researcher in order to highlight the field-dependent nature of careers and thus of indicators derived from them. The uniform-looking organizational careers of researchers hide the fact that these careers depend on conditions not immediately seen by an organization. This makes both the use of global standards (which are inevitably derived from one or few 'blueprint fields') and the aggregation of career indicators across fields dangerous. Researchers' careers are as different as the knowledge they produce.

References

Barley, Stephen R., 1989.

Careers, identities, and institutions: The legacy of the Chicago School of Sociology. Michael B. Arthur, Douglas T. Hall and Barbara S. Lawrence (eds.), Handbook of Career Theory. New York: Cambridge University Press, 41-66.

Bornmann, L. and D. Daniel, 2009.

The state of h index research. EMBO reports 10, 2 - 6. Butler, Linda, and Martijn S. Visser, 2006. Extending citation analysis to non-source items. Scientometrics 66: 327-343.

Clark, Burton R., 1983.

The higher education system: Academic Organization in Cross-National Perspective. Berkeley: University of California Press.

Costas, R. & Bordons, M. (2005).

Bibliometric indicators at the micro-level: some results in the area of natural resources at the Spanish CSIC. Research Evaluation, 14(2), 110-120.

Costas, R.; van Leeuwen, T.N. & Bordons, M.,2010.

A bibliometric classificatory approach for the study and assessment of research performance at the individual level: the effects of age on productivity and impact. Journal of the American Society for Information Science and Technology, 61(8), 1564-1581.

Dalton, G.W., P.H. Thompson, and R.L. Price, 1977.

The four stages of professional careers - a new look at performance by professionals. Organizational Dynamics 6: 19-42.

Gläser, Jochen, 2001.

Macrostructures, careers and knowledge production: a n

Gläser, Jochen, 2006.

Wissenschaftliche Produktionsgemeinschaften. Die soziale Ordnung der Forschung. Frankfurt am Main: Campus.

Sørensen, Aage B., 1992.

Wissenschaftliche Werdegänge und akademische Arbeitsmärkte. Karl-Ulrich Mayer (ed.), Generationsdynamik in der Forschung. Frankfurt a. M.: Campus, 83-121.

Gläser, Jochen, and Grit Laudel, 2009.

Identifying individual research trails. In Proceedings of the 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, 14-17 July 2009.

Hirsch, J.E., 2005.

An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America 102 (46), 16569-16572.

Laudel, Grit, 2005a.

Migration Currents among the scientific elite. Minerva 43: 377-395.

Laudel, Grit, 2005b.

Is external funding a valid indicator for research performance? Research Evaluation 14: 27-34. Research Assessment Exercise (RAE), 2008. RAE 2008 quality profiles, UOA 9 Psychiatry, Neuroscience and Clinical Psychology. www.rae.ac.uk/results/qualityProfile.aspx?id=9&type=uoa (accessed 11 April 2011)

Van Raan, Anthony F. J., 1996.

Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. Scientometrics 36: 397-420.

van Raan, Anthony J. F., 2000.

The Pandora's Box of Citation Analysis: Measuring Scientific Excellence - The Last Evil? Blaise Cronin and Helen Barsky Atkins (eds.), The Web of Knowledge. Medford, New Jersey: Information Today, Inc, 301-319.

Authors

Grit Laudel
University of Twente, Center for Higher Education Policy Studies,
P.O. Box 217, NL-7500AE Enschede
MAIL g.laudel@utwente.nl

Jochen Gläser
TU Berlin, Center for Technology and Society,
Hardenbergstraße 16-18,D-10623 Berlin,
MAIL Jochen.Glaser@ztg.tu-berlin.de